On a question of Mordell

Andrew R. Booker[a,1] **and Andrew V. Sutherland**[b,1]

[a]School of Mathematics, University of Bristol, Bristol BS8 1UG, United Kingdom; and [b]Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139

Edited by Kenneth A. Ribet, University of California, Berkeley, CA, and approved January 27, 2021 (received for review October 26, 2020)

We make several improvements to methods for finding integer solutions to $x^3+y^3+z^3=k$ for small values of k. We implemented these improvements on Charity Engine's global compute grid of 500,000 volunteer PCs and found new representations for several values of k, including 3 and 42. This completes the search begun by Miller and Woollett in 1954 and resolves a challenge posed by Mordell in 1953.

Diophantine equations | Hilbert's tenth problem | sums of three cubes

> "I think the problem, to be quite honest with you, is that you've never actually known what the question is." Douglas Adams, *The Hitchhiker's Guide to the Galaxy*, p. 183

1. Introduction

Let k be an integer with $k \not\equiv \pm 4 \pmod 9$. Heath-Brown (1) has conjectured that there are infinitely many triples $(x, y, z) \in \mathbb{Z}^3$ such that

$$x^3+y^3+z^3=k. \qquad [1.1]$$

Interest in this Diophantine equation goes back at least to Mordell (2), who asked whether there are any solutions to Eq. **1.1** for $k=3$ other than permutations of $(1,1,1)$ and $(4,4,-5)$. The following year, Miller and Woollett (3) used the electronic delay storage automatic calculator at Cambridge to run the first in a long line of computer searches attempting to answer Mordell's question, and also expanded the search to all positive $k \le 100$.

In this paper, we build on the approach of the first author in ref. 4, and find the following new solutions to Eq. **1.1**:

$$\begin{aligned}
569{,}936{,}821{,}221{,}962{,}380{,}720^3+(-569{,}936{,}821{,}113{,}563{,}493{,}509)^3+(-472{,}715{,}493{,}453{,}327{,}032)^3&=3,\\
(-80{,}538{,}738{,}812{,}075{,}974)^3+80{,}435{,}758{,}145{,}817{,}515^3+12{,}602{,}123{,}297{,}335{,}631^3&=42,\\
(-385{,}495{,}523{,}231{,}271{,}884)^3+383{,}344{,}975{,}542{,}639{,}445^3+98{,}422{,}560{,}467{,}622{,}814^3&=165,\\
143{,}075{,}750{,}505{,}019{,}222{,}645^3+(-143{,}070{,}303{,}858{,}622{,}169{,}975)^3+(-6{,}941{,}531{,}883{,}806{,}363{,}291)^3&=579,\\
(-74{,}924{,}259{,}395{,}610{,}397)^3+72{,}054{,}089{,}679{,}353{,}378^3+35{,}961{,}979{,}615{,}356{,}503^3&=906.
\end{aligned}$$

In particular, we answer Mordell's question and complete Miller and Woollett's search by finding at least one solution to Eq. **1.1** for all $k \le 100$ for which there are no local obstructions.

The algorithm used in ref. 4 is a refinement of an approach originally suggested in ref. 5, which is based on the following observation. Let us first assume that $|x|>|y|>|z|$ and define

$$d := |x+y|.$$

Then d is nonzero, and the solutions to Eq. **1.1** are precisely the triples (x, y, z) for which z is a cube root of k modulo d and the integer

$$\Delta(d,z) := 3d(4|k-z^3|-d^3) \quad \text{is a perfect square.} \qquad [1.2]$$

Solutions that do not satisfy $|x|>|y|>|z|$ can be efficiently found by other means: After a suitable permutation, either $x=-y$ and z is a cube root of k, or $y=z$ and we seek a solution to the Thue equation $x^3+2y^3=k$, which can be easily handled. If $|x|>|y|>|z|$ and we also assume $|z|>\sqrt{k}$, then we must have $0<d<\alpha|z|$, where $\alpha=\sqrt[3]{2}-1\approx 0.25992$; see ref. 4, section 2 for details. Solutions with $|z|\le\sqrt{k}$ are easily found by solving $x^3+y^3=k-z^3$ for each fixed z with $|z|\le\sqrt{k}$.

Significance

A Diophantine equation is a polynomial equation to which one seeks solutions in integers. There is a notable disparity between the difficulty of stating Diophantine equations and that of solving them. This feature was formalized in the 20th century by Matiyasevich's negative answer to Hilbert's tenth problem: It is impossible to tell whether some Diophantine equations have solutions or not. One need not look very far to find examples whose status is unknown. A striking example was noted by Mordell in 1953: The equation $x^3+y^3+z^3=3$ has the solutions (1, 1, 1) and (4, 4, −5) (and permutations); are there any others? This paper concludes a 65-y search with an affirmative answer to Mordell's question and strongly supports a related conjecture of Heath-Brown.

A.R.B. and A.V.S. performed research and wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

See online for related content such as Commentaries.

[1]To whom correspondence may be addressed. Email: Andrew.Booker@bristol.ac.uk or drew@math.mit.edu.

Published March 10, 2021.

This leads to an algorithm that searches for solutions with $|z| \le B$ by enumerating positive integers $d \le \alpha B$, and, for each such d, determining the residue classes of all cube roots of k modulo d and searching the corresponding arithmetic progressions for values of $z \in [-B, B]$ that make $\Delta(d, z)$ a square. With suitable optimizations, including sieving arithmetic progressions to quickly rule out integers that are not squares modulo primes in a suitably chosen set, this leads to an algorithm that requires only $O\left(B(\log\log B)(\log\log\log B)\right)$ operations on integers in $[0, B]$ for any fixed value of k. An attractive feature of this algorithm is that it finds all solutions with $\min\{|x|, |y|, |z|\} \le B$, even those for which $\max\{|x|, |y|, |z|\}$ may be much larger than B (note that this is the case in our solution for $k = 3$).

This algorithm was used in ref. 4 to find solutions for $k = 33$ and $k = 795$, leaving only the following 11 $k \le 1{,}000$ unresolved:

$$42,\ 114,\ 165,\ 390,\ 579,\ 627,\ 633,\ 732,\ 906,\ 921,\ 975. \qquad \textbf{[1.3]}$$

The search in ref. 4 also ruled out any solutions for these k with $\min\{|x|, |y|, |z|\} \le 10^{16}$.

Here we make several improvements to this method in ref. 4 that allow us to find a new solution for $k = 3$ as well as four of the outstanding k listed above.

- Instead of a single parameter B bounding $|z| \le B$ and $0 < d \le \alpha B$, we use independent bounds $d_{\max}$ on d and $z_{\max}$ on $|z|$, whose ratio we optimize via an analysis of the expected distribution of $|z|/d$; this typically leads to a $z_{\max}/d_{\max}$ ratio that is 10 to 20 times larger than the ratio $1/\alpha \approx 3.847332$ used in ref. 4.
- Rather than explicitly representing a potentially large set of sieved arithmetic progressions containing candidate values of z for a given d, we implicitly represent them as intersections of arithmetic progressions modulo the prime power factors of d and auxiliary primes. This both improves the running time and reduces the memory footprint of the algorithm, allowing for much larger values of $|z|$.
- We dynamically optimize the choice of auxiliary primes used for sieving based on the values of k and d; when d is much smaller than $z_{\max}$, this can reduce the number of candidate values of z by several orders of magnitude.
- We exploit 3-adic and cubic reciprocity constraints for all $k \equiv \pm 3 \pmod 9$; for the values of k listed in Eq. **1.3**, this reduces the average number of z we need to check for a given value of d by a factor of between 2 and 4 compared to the congruence constraints used in ref. 4, which did not use cubic reciprocity for $k \ne 3$.

Along the way, we compute, to high precision, the expected density of solutions to Eq. **1.1** conjectured by Heath-Brown (1), and compare it with the numerical data compiled by Huisman (6) for $k \in [3, 1{,}000]$ and $\max\{|x|, |y|, |z|\} \le 10^{15}$. The data strongly support Heath-Brown's conjecture that Eq. **1.1** has infinitely many solutions for all $k \not\equiv \pm 4 \pmod 9$.

2. Density Computations

In this section, we study Heath-Brown's conjecture in detail. In particular, we explain how to compute the conjectured density of solutions to high precision and compare the results with available numerical data. We further study the densities of divisors $d \mid z^3 - k$ and arithmetic progressions $z \pmod d$ that occur in our algorithm, which informs the choice of parameters used in our computations.

Let k be a cube-free integer with $k \ge 3$ and $k \not\equiv \pm 4 \pmod 9$. Define $K = \mathbb{Q}(\sqrt[3]{k})$ and $F = \mathbb{Q}(\sqrt{-3})$, and let $\mathfrak{o}_K$ and $\mathfrak{o}_F$ be their respective rings of integers. We have $\mathfrak{o}_F = \mathbb{Z}[\zeta_6]$, where $\zeta_6 = \frac{1+\sqrt{-3}}{2}$ is a generator of the unit group $\mathfrak{o}_F^\times$. Also, $\mathrm{Disc}(F) = -3$ and $\mathrm{Disc}(K) = -3f^2$, where, by ref. 7, lemma 2.1,

$$f = \Big(\prod_{p \mid k} p\Big) \cdot \begin{cases} 1 & \text{if } k \equiv \pm 1 \pmod 9, \\ 3 & \text{otherwise.} \end{cases}$$

We define two modular forms related to F and K. First, let f_1 be the modular form of weight 1 and level $|\mathrm{Disc}(K)|$ such that $\zeta_K(s) = \zeta(s)L(s, f_1)$. It follows, from the ramification description in ref. 7, section 2.1, that rational primes p decompose into prime ideals of $\mathfrak{o}_K$ as follows (subscripts denote inertia degrees):

$$p\mathfrak{o}_K = \begin{cases} \mathfrak{p}_1\mathfrak{p}_2 & \text{if } p \equiv 2 \pmod 3 \text{ and } p \nmid k, \\ \mathfrak{p}_1\mathfrak{p}_1'\mathfrak{p}_1'' & \text{if } p \equiv 1 \pmod 3 \text{ and } p \nmid k \text{ and } k \text{ is a cube modulo } p, \\ \mathfrak{p}_3 & \text{if } p \equiv 1 \pmod 3 \text{ and } p \nmid k \text{ and } k \text{ is not a cube modulo } p, \\ \mathfrak{p}_1^2\mathfrak{p}_1' & \text{if } p = 3 \text{ and } k \equiv \pm 1 \pmod 9, \\ \mathfrak{p}_1^3 & \text{otherwise.} \end{cases}$$

From this data, we find that the local Euler factor of $L(s, f_1)$ at p is

$$L_p(s, f_1) = \frac{1}{1 - c_p(k)p^{-s} + \left(\frac{\mathrm{Disc}(K)}{p}\right)p^{-2s}}, \quad \text{where } c_p(k) := \begin{cases} 2 & \text{if } p \nmid k,\ p \equiv 1 \pmod 3 \text{ and } k^{(p-1)/3} \equiv 1 \pmod p, \\ -1 & \text{if } p \nmid k,\ p \equiv 1 \pmod 3 \text{ and } k^{(p-1)/3} \not\equiv 1 \pmod p, \\ 1 & \text{if } p = 3 \text{ and } k \equiv \pm 1 \pmod 9, \\ 0 & \text{otherwise.} \end{cases}$$

Now let $\sigma : F \to \mathbb{C}$ be the unique embedding for which we have $\Im\sigma(\sqrt{-3}) > 0$. Let $\chi_f : (\mathfrak{o}_F/3\mathfrak{o}_F)^\times \to \mathbb{C}^\times$ be the character defined by $\chi_f(\zeta_6 + 3\mathfrak{o}_F) = \sigma(\zeta_6^{-1})$, and define $\chi_\infty(z) := \sigma(z)/|\sigma(z)|$. Let χ be the Grössencharakter of F defined by

$$\chi(\alpha\mathfrak{o}_F) := \begin{cases} \chi_\infty(\alpha)\chi_f(\alpha + 3\mathfrak{o}_F) & \text{if } \alpha \in \mathfrak{o}_F \setminus \sqrt{-3}\mathfrak{o}_F, \\ 0 & \text{if } \alpha \in \sqrt{-3}\mathfrak{o}_F. \end{cases}$$

By automorphic induction, there is a holomorphic newform f_2 of weight 2 and level $|\operatorname{Disc}(F)|N(3\mathfrak{o}_F)=27$ such that $L(s,f_2)=L(s-\frac{1}{2},\chi)$.

Given a prime $p\equiv 1 \pmod 3$, let a_p denote the unique integer for which $a_p\equiv 1 \pmod 3$ and $4p=a_p^2+27b^2$ for some $b\in\mathbb{Z}_{>0}$. Let $\alpha=\frac{a_p+3b\sqrt{-3}}{2}$ and $\mathfrak{p}=\alpha\mathfrak{o}_F$, so that $p\mathfrak{o}_F=\mathfrak{p}\bar{\mathfrak{p}}$. We have

$$\alpha+1=\frac{a_p+2+3b\sqrt{-3}}{2}=3\left(\frac{\frac{a_p+2}{3}+b\sqrt{-3}}{2}\right)\in 3\mathfrak{o}_F,$$

so $\chi_f(\alpha)=\chi_f(-1)=-1$. Thus, $\chi(\mathfrak{p})=-\frac{\sigma(\alpha)}{\sqrt{p}}$ and $\chi(\bar{\mathfrak{p}})=-\frac{\overline{\sigma(\alpha)}}{\sqrt{p}}$, so the Euler factor of $L(s,f_2)$ at p (in its arithmetic normalization) is

$$\frac{1}{(1-\chi(\mathfrak{p})p^{\frac{1}{2}-s})(1-\chi(\bar{\mathfrak{p}})p^{\frac{1}{2}-s})}=\frac{1}{1+a_pp^{-s}+p^{1-2s}}.$$

For a prime $p\equiv 2 \pmod 3$, we have $\chi(p\mathfrak{o}_F)=\chi_\infty(p)\chi_f(p)=\chi_f(-1)=-1$, so the corresponding Euler factor is

$$\frac{1}{1-\chi(p\mathfrak{o}_F)N(p\mathfrak{o}_F)^{\frac{1}{2}-s}}=\frac{1}{1+p^{1-2s}}.$$

Finally, $\chi(\sqrt{-3})=0$, so the Euler factor at $p=3$ is 1.

In summary, if we extend the definition of a_p so that $a_p=0$ for $p\not\equiv 1 \pmod 3$, then the Euler factor of $L(s,f_2)$ at p is

$$L_p(s,f_2)=\frac{1}{1+a_pp^{-s}+\left(\frac{9}{p}\right)p^{1-2s}}.$$

A. Solution Density. Define

$$\sigma_p:=\lim_{e\to\infty}\frac{\#\{(x,y,z) \pmod{p^e}: x^3+y^3+z^3\equiv k \pmod{p^e}\}}{p^{2e}}.$$

Then, as calculated by Heath-Brown (1), we have

$$\sigma_p=\begin{cases}1+\frac{3c_p(k)}{p}-\frac{a_p}{p^2} & \text{if } p\nmid 3k,\\ 1+\frac{(p-1)a_p-1}{p^2} & \text{if } p\mid k \text{ and } p\neq 3,\\ \frac{1}{3}\#\{(x,y,z) \pmod 3: x^3+y^3+z^3\equiv k \pmod 9\} & \text{if } p=3.\end{cases}$$

Now let $h\colon\mathbb{R}^3\to\mathbb{R}_{\geq 0}$ be a height function, by which we mean a function that is continuous, symmetric in its inputs, and satisfies

- $h(x,y,z)>0$ when $x^3+y^3+z^3=0$ and $xyz\neq 0$,
- $h(\lambda x,\lambda y,\lambda z)=|\lambda|h(x,y,z)$ for any $\lambda\in\mathbb{R}^\times$.

The real density of solutions to $x^3+y^3+z^3=0$ with height in the interval $\mathcal{H}:=[H_1,H_2]$ can then be computed as follows.

For $\varepsilon>0$, define

$$S(\varepsilon):=\left\{(x,y,z)\in\mathbb{R}^3: h(x,y,z)\in\mathcal{H},\ |x^3+y^3+z^3|\leq\varepsilon\right\},$$
$$T(\varepsilon):=\left\{(x,y,z)\in\mathbb{R}^3: x\geq y\geq z\geq 0,\ h(x,-y,-z)\in\mathcal{H},\ |x^3-y^3-z^3|\leq\varepsilon\right\},$$

so that $\operatorname{vol}(S(\varepsilon))/\operatorname{vol}(T(\varepsilon))\to 12$ as $\varepsilon\to 0^+$. We may then compute

$$\begin{aligned}\lim_{\varepsilon\to 0^+}(2\varepsilon)^{-1}\operatorname{vol}(S(\varepsilon))&=12\lim_{\varepsilon\to 0^+}(2\varepsilon)^{-1}\operatorname{vol}(T(\varepsilon))\\&=12\int_0^\infty\int_z^\infty \mathbf{1}_{h(\sqrt[3]{y^3+z^3},-y,-z)\in\mathcal{H}}\frac{dy}{3(y^3+z^3)^{2/3}}\,dz\\&=4\int_0^\infty\int_1^\infty \mathbf{1}_{zh(\sqrt[3]{t^3+1},-t,-1)\in\mathcal{H}}\frac{dt}{(t^3+1)^{2/3}}\frac{dz}{z}\\&=4\int_1^\infty\int_{H_1/h(\sqrt[3]{t^3+1},-t,-1)}^{H_2/h(\sqrt[3]{t^3+1},-t,-1)}\frac{dz}{z}\frac{dt}{(t^3+1)^{2/3}}=\sigma_\infty\log\frac{H_2}{H_1},\end{aligned}\qquad [2.1]$$

where $\sigma_\infty=4\int_1^\infty(t^3+1)^{-2/3}\,dt=\frac{2}{3}\frac{\Gamma(\frac{1}{3})^2}{\Gamma(\frac{2}{3})}$.

Heath-Brown conjectures that the number $n(B)$ of solutions to Eq. **1.1**, up to permutation, satisfying $\max\{|x|,|y|,|z|\}\leq B$ is asymptotic to

$$e(B):=\rho_{\text{sol}}\log B\quad\text{as } B\to\infty,\qquad\text{where } \rho_{\text{sol}}:=\tfrac{1}{6}\sigma_\infty\prod_p\sigma_p.$$

As shown above, the real density does not depend on the precise choice of the height function h. We thus conjecture that the same asymptotic density applies to the solutions satisfying $h(x,y,z)\leq B$ for any similar choice of h, including, for example,

$$\min\{|x|,|y|,|z|\},\quad |xyz|^{\frac{1}{3}},\quad\text{and}\quad d=\min\{|x+y|,|x+z|,|y+z|\}.$$

Let us now define

$$r_p := \frac{\sigma_p}{L_p(1,f_1)^3 L_p(2,f_2) L_p(2,f_1)^{-6} \zeta_p(2)^{-6} L_p(2,(\frac{\cdot}{3}))^{-3}}, \quad \text{where } \zeta_p(s) = \frac{1}{1-p^{-s}} \text{ and } L_p(s,(\tfrac{\cdot}{3})) = \frac{1}{1-\left(\frac{p}{3}\right)p^{-s}}.$$

A straightforward calculation shows that

$$r_p = 1 - \frac{3a_p c_p(k) + O(1)}{p^3}.$$

Since $-a_p c_p(k)$ is the coefficient of p^{-s} in the Rankin–Selberg L-function $L(s, f_1 \boxtimes f_2)$, we expect square-root cancellation in the product $\prod_p r_p$. Under the generalized Riemann hypothesis (GRH), for large X, we have

$$\prod_p \sigma_p = \big(1 + O(X^{-2}\log X)\big) L(1,f_1)^3 L(2,f_2) L(2,f_1)^{-6} \zeta(2)^{-6} L(2,(\tfrac{\cdot}{3}))^{-3} \prod_{p \le X} r_p. \qquad \textbf{[2.2]}$$

Applying Eq. **2.2** with $X = 10^9$ allows us to compute the solution densities ρ_{sol} to roughly 18 digits of precision for all cube-free $k \le 1{,}000$. To evaluate the L-functions, we used the extensive functionality available for that purpose in PARI/GP (8). Since our goal is merely to gather some statistics, we content ourselves with a heuristic estimate of the error in this computation, although it could be rigorously certified with more work. Some examples are shown in Table 1.

We compared Huisman's dataset to an average form of Heath-Brown's conjecture as follows. For an integer $K \ge 3$, define

$$N_K(B) := \#\{(k,x,y,z) \in \mathbb{Z}^4 : x^3 + y^3 + z^3 = k \text{ cube-free}, \ 3 \le k \le K, \ |z| \le |y| \le |x| \le B\} \quad \text{and} \quad \rho_K := \sum_{\substack{k \in \mathbb{Z} \cap [3,K] \\ k \text{ cube-free}}} \rho_{\mathrm{sol}}(k).$$

Then Heath-Brown's conjecture implies that, for fixed K, we have $N_K(B) \sim \rho_K \log B$ as $B \to \infty$. The plot in Fig. 1 compares $N_{1000}(B)$ for $B \in [10^{7.5}, 10^{15}]$, computed from Huisman's (6) data, with $\rho_{1000}\log B + C$, where $\rho_{1000} \approx 363.869$ and $C \approx -679.4$ was chosen to minimize the mean square difference. Out of 6,256 points, the two plots never differ by more than 42, which gives strong evidence for Heath-Brown's conjecture, at least on average over k.

B. Divisor and Arithmetic Progression Densities. In this section, we assume that $k \equiv \pm 3 \pmod 9$ and derive estimates for the density of arithmetic progressions arising from cube roots of k modulo d. Define

$$\delta_d := \begin{cases} 1 & \text{if } \exists z \in \mathbb{Z} \text{ s.t. } z^3 \equiv k \pmod d \text{ and } \mathrm{ord}_p(d) \in \{0, \mathrm{ord}_p(k/3)\} \ \forall p \mid k, \\ 0 & \text{otherwise,} \end{cases} \quad \text{and} \quad F(s) := \sum_{d=1}^{\infty} \frac{\delta_d}{d^s}.$$

As shown in ref. 4, any d arising from a solution to Eq. **1.1** must satisfy $\delta_d = 1$, and we only consider such d in our algorithm.

Table 1. Selected ρ_{sol} and $\lceil \exp(1/\rho_{\mathrm{sol}}) \rceil = \min\{B \in \mathbb{Z} : e(B) \ge 1\}$ for $k \le 1{,}000$, including 10 smallest ρ_{sol} and all k with $n(10^{15}) = 0$

			$B = 10^5$		$B = 10^{10}$		$B = 10^{15}$	
k	ρ_{sol}	$\lceil \exp(1/\rho_{\mathrm{sol}}) \rceil$	$e(B)$	$n(B)$	$e(B)$	$n(B)$	$e(B)$	$n(B)$
858	0.028504	1,723,846,985,902,459	0.328	1	0.656	2	0.984	2
276	0.031854	43,031,002,119,138	0.367	1	0.733	1	1.100	2
390	0.032935	15,358,736,844,736	0.379	0	0.758	0	1.138	0
516	0.033062	13,665,771,588,173	0.381	0	0.761	1	1.142	1
663	0.033196	12,097,471,969,974	0.382	0	0.764	1	1.147	1
975	0.038722	164,297,126,902	0.446	0	0.892	0	1.337	0
165	0.039636	90,602,378,809	0.456	0	0.913	0	1.369	0
555	0.042706	14,770,444,441	0.492	1	0.983	2	1.475	2
921	0.044142	6,895,540,744	0.508	0	1.016	0	1.525	0
348	0.044632	5,378,175,303	0.514	2	1.028	2	1.542	3
906	0.049745	537,442,063	0.573	0	1.145	0	1.718	0
579	0.050838	348,939,959	0.585	0	1.171	0	1.756	0
114	0.058459	26,853,609	0.673	0	1.346	0	2.019	0
3	0.061052	12,985,612	0.703	2	1.406	2	2.109	2
732	0.063137	7,561,540	0.727	0	1.454	0	2.181	0
633	0.079660	283,059	0.917	0	1.834	0	2.751	0
33	0.088833	77,422	1.023	0	2.045	0	3.068	0
795	0.089491	71,273	1.030	0	2.061	0	3.091	0
42	0.113449	6,732	1.306	0	2.612	0	3.918	0
627	0.129565	2,249	1.492	0	2.983	0	4.475	0

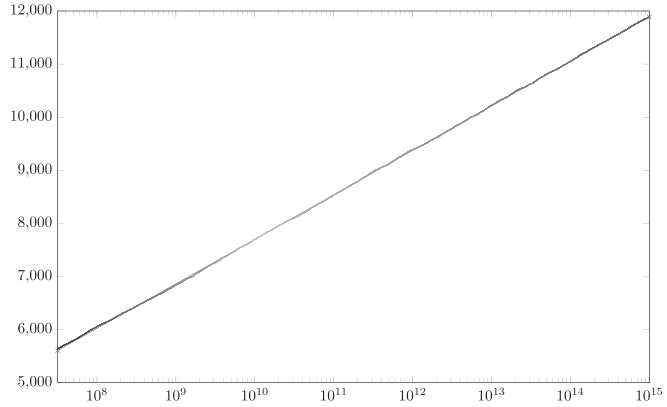

Fig. 1. Scatter plot of $N_{1000}(B)$ as a function of $B \in [10^{7.5}, 10^{15}]$ based on Huisman's dataset (6), compared to the line $\rho_{1000} \log B + C$.

For $p \nmid k$ and $e > 0$, we have $\delta_{p^e} = \frac{c_p(k)+2-\left(\frac{p}{3}\right)}{3}$, so that $F(s) = \prod_p F_p(s)$, where

$$F_p(s) := \begin{cases} \left(1 - \frac{c_p(k)+2-\left(\frac{p}{3}\right)}{3p^s}\right)^{-1} & \text{if } p \nmid k, \\ 1 + p^{-\operatorname{ord}_p(k)s} & \text{if } p \mid \frac{k}{3}, \\ 1 & \text{if } p = 3. \end{cases}$$

For $p \nmid k$, the local factor $\frac{F_p(s)^3 L_p(s,(\frac{\cdot}{3}))}{\zeta_p(s)^2 L_p(s,f_1)}$ is $1 + O(p^{-3s})$. Therefore, $F(s)^3$ has meromorphic continuation to $\Re(s) > \frac{1}{3}$, with a pole of order 2 at $s = 1$ and no other poles in the region $\{s \in \mathbb{C} : \Re(s) \geq 1\}$. By ref. 9, theorem 3.1, it follows that

$$\sum_{d \leq d_{\max}} \delta_d \sim \rho_{\text{div}} \frac{d_{\max}}{\sqrt[3]{\log d_{\max}}} \quad \text{as } d_{\max} \to \infty, \qquad \text{where } \rho_{\text{div}} := \frac{\left(\lim_{s\to 1} F(s)^3 (s-1)^2\right)^{\frac{1}{3}}}{\Gamma\left(\frac{2}{3}\right)}.$$

In turn, we have

$$\lim_{s\to 1} F(s)^3 (s-1)^2 = \left(1 + O(X^{-2})\right) \frac{L(1,f_1)}{L\left(1,\left(\frac{\cdot}{3}\right)\right)} \prod_{p \leq X} \frac{F_p(1)^3 L_p\left(1,\left(\frac{\cdot}{3}\right)\right)}{\zeta_p(1)^2 L_p(1,f_1)}.$$

Let us now define

$$G(s) := \sum_{d=1}^{\infty} \frac{\delta_d r_d(k)}{d^s}, \quad \text{where } r_d(k) = \#\{z \pmod d : z^3 \equiv k \pmod d\}.$$

Then $G(s) = \prod_p G_p(s)$, where

$$G_p(s) := \begin{cases} 1 + \frac{1+c_p(k)}{p^s - 1} & \text{if } p \nmid k, \\ 1 + p^{(1-s)\operatorname{ord}_p(k) - 1} & \text{if } p \mid \frac{k}{3}, \\ 1 & \text{if } p = 3. \end{cases}$$

For $p \nmid k$, we have $\frac{G_p(s) L_p(2s,f_1) \zeta_p(2s)}{L_p(s,f_1)\zeta_p(s)} = 1 + O(p^{-3s})$. Therefore,

$$\sum_{d \leq d_{\max}} \delta_d r_d(k) \sim \rho_{\text{ap}} d_{\max} \quad \text{as } d_{\max} \to \infty, \qquad \text{where } \rho_{\text{ap}} := \lim_{s\to 1} G(s)(s-1).$$

In turn, we have

$$\rho_{\text{ap}} = \left(1 + O(X^{-2})\right) \frac{L(1,f_1)}{L(2,f_1)\zeta(2)} \prod_{p \leq X} \frac{G_p(1) L_p(2,f_1) \zeta_p(2)}{L_p(1,f_1)\zeta_p(1)}.$$

Table 2 lists estimates $\rho_{\text{ap}} d_{\max}$ for the number $\pi_{\text{ap}}(d_{\max})$ of arithmetic progressions modulo $d \leq d_{\max}$, as well as estimates $\rho_{\text{div}} d_{\max} / \sqrt[3]{\log d_{\max}}$ for the number $\pi_{\text{div}}(d_{\max})$ of admissible $d \leq d_{\max}$, along with the ratios of these quantities.

Remark 2.1: The average number of arithmetic progressions modulo $d \leq d_{\max}$ listed in Table 2 is strikingly small. Even for $d_{\max} = 10^{24}$, which is well beyond the feasible range, the average is around 3 and never above 3.5 for any of the listed k.

Remark 2.2: For any fixed choice of the ratio $R = z_{\max}/d_{\max}$, the total running time of our algorithm is roughly proportional to $\rho_{\text{ap}} d_{\max}$. The constant of proportionality can be estimated by running the algorithm on a suitable sample of $d \leq d_{\max}$. These estimates allow us to efficiently manage resource allocation in large distributed computations; see section 5 for details.

Table 2. Comparison of estimated and actual counts of arithmetic progressions modulo $d \le d_{\max} = 10^{12}$ for various k of interest

k	$\rho_{\mathrm{ap}} d_{\max}$	$\pi_{\mathrm{ap}}(d_{\max})$	$\frac{\rho_{\mathrm{div}} d_{\max}}{\sqrt[3]{\log d_{\max}}}$	$\pi_{\mathrm{div}}(d_{\max})$	$\frac{\rho_{\mathrm{ap}} \sqrt[3]{\log d_{\max}}}{\rho_{\mathrm{div}}}$	$\frac{\pi_{\mathrm{ap}}(d_{\max})}{\pi_{\mathrm{div}}(d_{\max})}$
3	476,709,085,641	476,709,082,386	221,480,415,360	222,316,170,600	2.152	2.144
42	439,262,042,312	439,262,055,314	194,525,166,395	195,043,114,314	2.258	2.252
114	346,031,225,026	346,031,232,985	169,944,552,313	169,697,769,695	2.036	2.039
165	398,768,628,911	398,768,635,237	201,820,401,130	201,648,107,384	1.976	1.978
390	361,424,697,190	361,424,750,258	170,411,108,873	170,119,932,464	2.121	2.125
579	467,532,879,762	467,532,936,236	220,746,986,113	221,627,128,720	2.118	2.110
627	544,308,148,137	544,308,117,802	238,234,806,279	240,026,258,762	2.285	2.268
633	510,771,397,972	510,771,391,669	227,368,579,096	228,697,959,163	2.246	2.233
732	396,862,883,895	396,862,943,789	145,013,347,786	145,167,910,326	2.737	2.734
906	353,110,285,004	353,110,236,539	166,128,603,588	165,813,813,631	2.126	2.130
921	420,143,131,383	420,143,101,621	212,693,499,876	212,924,474,063	1.975	1.973
975	461,977,372,770	461,977,396,756	194,140,103,965	194,481,735,572	2.380	2.375

3. Cubic Reciprocity

In ref. 10, Cassels used cubic reciprocity to prove that, whenever $x, y, z \in \mathbb{Z}$ satisfy $x^3 + y^3 + z^3 = 3$, we must have $x \equiv y \equiv z \pmod 9$. For fixed $d = |x + y|$, it follows that z is determined modulo 81. Colliot-Thélène and Wittenberg (11) later recast this phenomenon in terms of Brauer–Manin obstructions, and showed that, for any k, the solutions to Eq. **1.1** are always forbidden for some residue classes globally but not locally.* In this section, we extend Cassels' analysis to all cube-free $k \equiv \pm 3 \pmod 9$, and derive constraints on the residue class of $z \pmod q$ for a certain modulus $q \mid 27k$. We assume throughout that $k \equiv 3\epsilon \pmod 9$ for a fixed $\epsilon \in \{\pm 1\}$.

Given $\alpha, \beta \in \mathfrak{o}_F$ with $\beta \notin \sqrt{-3}\mathfrak{o}_F$, let $\left(\frac{\alpha}{\beta}\right)_3$ be the cubic residue symbol, as defined in ref. 12, chapters 9 and 14. Put $\zeta_3 = \frac{-1+\sqrt{-3}}{2} \in \mathfrak{o}_F$. For integers x, y satisfying $x \equiv y \equiv \epsilon \pmod 3$, define

$$\chi_k(x, y) := \zeta_3^{\epsilon(y-x)/3} \left(\frac{\zeta_3 x + \zeta_3^{-1} y}{k/3} \right)_3.$$

Note that $\chi_k(x, y)$ depends only on the residue classes of $x, y \pmod{3k}$.

Definition 3.1: We say that a pair $(d, z) \in \mathbb{Z}^2$ is admissible if there exist $x, y \in \mathbb{Z}$ satisfying the following conditions:

1) $x + y \equiv -\epsilon \left(\frac{d}{3}\right) d \pmod{27k}$;
2) $x^3 + y^3 + z^3 \equiv k \pmod{81k}$;
3) $\{\chi_k(x, y), \chi_k(x, z), \chi_k(y, z)\} \subseteq \{0, 1\}$.

Note that this definition depends only on the residue classes of $d, z \pmod{27k}$.

Lemma 3.2. *Let $(x, y, z) \in \mathbb{Z}^3$ be a solution to* Eq. **1.1**, *and let $d := |x + y|$. Then (d, z) is admissible.*

Proof: Recall that $k \equiv 3\epsilon \pmod 9$. Since every cube is congruent to 0 or $\pm 1 \pmod 9$, we have $x \equiv y \equiv z \equiv \epsilon \pmod 3$, so that $x + y \equiv -\epsilon \equiv -\epsilon \left(\frac{d}{3}\right) d \pmod 3$. As $d = |x + y|$, it follows that $x + y = -\epsilon \left(\frac{d}{3}\right) d$, so condition 1 of the definition is satisfied. Condition 2 then follows directly from Eq. **1.1**.

Now let

$$\gamma := \epsilon(\zeta_3 x + \zeta_3^{-1} y) = -\epsilon y + \epsilon(x - y)\zeta_3.$$

By ref. 12, chapter 9, example 19, we have

$$\chi_k(x, y) = \zeta_3^{\epsilon(y-x)/3} \left(\frac{\zeta_3 x + \zeta_3^{-1} y}{k/3} \right)_3 = \left(\frac{3}{\gamma} \right)_3 \left(\frac{\epsilon\gamma}{k/3} \right)_3 = \left(\frac{-3\epsilon}{\gamma} \right)_3 \left(\frac{\gamma}{-\epsilon k/3} \right)_3,$$

where the last equality follows from the fact that $(\alpha/\beta)_3$ depends only on the ideal $\beta\mathfrak{o}_F$ and $(\pm\epsilon/\beta)_3 = ((\pm\epsilon)^3/\beta)_3 = 1$. By cubic reciprocity (ref. 12, chapter 14, theorem 1), this equals

$$\left(\frac{-3\epsilon}{\gamma} \right)_3 \left(\frac{-\epsilon k/3}{\gamma} \right)_3 = \left(\frac{k}{\gamma} \right)_3.$$

Noting that $x^3 + y^3 = (x + y)\gamma\overline{\gamma}$, we have $k \equiv z^3 \pmod{\gamma\mathfrak{o}_F}$, whence

$$\chi_k(x, y) = \left(\frac{z^3}{\gamma} \right)_3 \in \{0, 1\},$$

and, by symmetry, we also have $\chi_k(x, z), \chi_k(y, z) \in \{0, 1\}$; thus condition 3 holds as well. □

*Thus strong approximation fails for Eq. **1.1**, but this is never enough to forbid the existence of integer solutions outright, so there is no Brauer–Manin obstruction.

Lemma 3.3. *Let*

$$q := 27k \prod_{\substack{p \mid k \\ \operatorname{ord}_p(k)=2 \\ p=2 \text{ or } c_p(2)=-1}} p^{-1},$$

and let $d, z, z' \in \mathbb{Z}$ *satisfy* $z' \equiv z \pmod q$. *Then* (d, z) *is admissible iff* (d, z') *is admissible.*

Proof: Suppose that (d, z) is admissible. Let p be a prime divisor of $27k/q$, and consider $z' \equiv z \pmod{27k/p}$. By the Chinese remainder theorem, it suffices to show that (d, z') is admissible in this case.

Set $a = (z' - z)/p$, so that $z' = z + ap$. Let x, y be integers satisfying the conditions in *Definition 3.1*, and let $x' = x + bp$, $y' = y - bp$ for some $b \in 27kp^{-2}\mathbb{Z}$. Then

$$(x')^3 + (y')^3 + (z')^3 \equiv x^3 + y^3 + z^3 + 3p[az^2 + b(x^2 - y^2)] \equiv 3p[az^2 + b(x^2 - y^2)] \pmod{p^2}.$$

If $p \mid (x^2 - y^2)$ and $p \nmid (x + y)$, then we have $x \equiv y \pmod p$, $p > 2$ and $p \nmid x$, which means that $2x^3 + z^3 \equiv 0 \pmod p$ and $2 \equiv (-z/x)^3 \pmod p$ is a cubic residue mod p. But $p > 2$ implies $c_p(2) = -1$, meaning $p \equiv 1 \pmod 3$ and $2^{(p-1)/3} \not\equiv 1 \pmod p$, so 2 cannot be a cubic residue mod p, and we must have $p \nmid (x^2 - y^2)$ or $p \mid (x + y)$.

If $p \nmid (x^2 - y^2)$, then we may choose b so that $b(x^2 - y^2) \equiv -az^2 \pmod p$, while, if $p \mid (x + y)$, then $p \mid z$ and any choice of b suffices. It follows that

$$(x')^3 + (y')^3 + (z')^3 \equiv k \pmod{81k}.$$

Moreover, we have $\chi_k(x', y') = \chi_k(x, y)$, $\chi_k(x', z') = \chi_k(x, z)$, and $\chi_k(y', z') = \chi_k(y, z)$, by inspection. Thus (d, z') is admissible, as desired. □

Thus, the definition of admissibility factors through $\mathbb{Z}/27k\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z}$.

Example 3.4: The table below shows the ratio

$$\frac{\sum_{d \pmod{27k}} \#\{z \pmod q : (d, z) \text{ is admissible}\}}{\sum_{d \pmod{27k}} \#\{z \pmod q : \exists x \pmod q \text{ s.t. } x^3 + (d - x)^3 + z^3 \equiv k \pmod{3q}\}},$$

which is the average density of admissible residues $z \pmod q$ among all locally permitted residues, for a few k of interest.

k	3	33	42	114	633
Density	0.250	0.590	0.970	0.962	0.585

Although the improvement is modest for some k, those cases still benefit from imposing local constraints mod q, some of which were not used in ref. 4; in particular, passing from mod 9 solutions to mod 81 solutions reduces the density by a factor of $4/9$.

A. Algorithm. Let $k \equiv 3\epsilon \pmod 9$ be a positive integer, and, for each positive integer m, let

$$\mathcal{C}(m) := \{z + m\mathbb{Z} : z^3 \equiv k \pmod m\} \subseteq \mathbb{Z}/m\mathbb{Z}$$

denote the set of cube roots of k modulo m. Let P be the set of primes $p \nmid k$ for which $\#\mathcal{C}(p) > 0$; for $p \in P$, we then have $\#\mathcal{C}(p) = 3$ if $p \equiv 1 \pmod 3$, and $\#\mathcal{C}(p) = 1$ otherwise.

Let A be a set of small auxiliary primes $p \nmid k$ whose product exceeds $d_{\max} z_{\max}$; in practical computations, we may take A to be the primes $p < 256$ not dividing k. Let $s := \epsilon \left(\frac{d}{3}\right)$, so that any solution to Eq. **1.1** with $d = |x + y|$ has $\operatorname{sgn} z = s$, and, for positive integers d and primes $p \nmid dk$, define

$$\mathcal{S}_d(p) := \begin{cases} \{z + p\mathbb{Z} : 3d(4s(z^3 - k) - d^3) \equiv \square \pmod p\} & \text{if } p > 2, \\ \{k + d + 2\mathbb{Z}\} & \text{if } p = 2. \end{cases}$$

Finally, let $c_1 > c_0 > 1$ and $c_2 > 1$ denote integers that we will choose to optimize performance (typically, $c_0 \approx 4$, $c_1 \approx 50$, and $c_2 \approx 6$), and let q be the divisor of $27k$ defined in *Lemma 3.3*.

Algorithm 3.5: Given $k, d_{\max}, z_{\max} \in \mathbb{Z}_{>0}$ with $k \equiv 3\epsilon \pmod 9$, enumerate all pairs $(d, z) \in \mathbb{Z}^2$ for which there exist $(x, y, z) \in \mathbb{Z}^3$ satisfying Eq. **1.1** with $|x| > |y| > |z|$, $\sqrt{k} < |z| \le z_{\max}$, and $|x + y| = d \le d_{\max}$ as follows:

Recursively enumerate all positive integers $d_0 = p_1^{e_1} \cdots p_n^{e_n} \le d_{\max}$, where $p_1 > \cdots > p_n$ are primes in P and $e_i \in \mathbb{Z}_{>0}$. For each such d_0, do the following:

1) For each positive divisor d_1 of $k/3$ with $\gcd(d_1, k/d_1) = 1$, set $d := d_0 d_1$ and let $\mathcal{A}_d(q)$ be the set of $z + q\mathbb{Z}$ for which (d, z) is admissible.
2) Set $a := 1$, and, if $c_1 q d_0 < z_{\max}$, then order the $p \nmid d$ in A by $\log \#\mathcal{S}_d(p)/\log p$, and, while $c_0 q d_0 pa < z_{\max}$, replace a by pa, where p is the next prime in the ordering.
3) Let b be the product of c_2 primes $p \in A$ not dividing da, chosen using either the ordering computed in the previous step or a fixed order.
4) Set $m := d_0 qa$, and let $\mathcal{Z}(m)$ be the subset of $\mathbb{Z}/m\mathbb{Z}$ that is identified with

$$\mathcal{C}(p_1^{e_1}) \times \cdots \times \mathcal{C}(p_n^{e_n}) \times \mathcal{A}_d(q) \times \prod_{p \mid a} \mathcal{S}_d(p)$$

via the Chinese remainder theorem. Let

$$\mathcal{Z}(m, s, z_{\max}) := \{z \in \mathbb{Z} : z + m\mathbb{Z} \in \mathcal{Z}(m),\ \operatorname{sgn} z = s,\ \text{and } |z| \le z_{\max}\}.$$

For each $z \in \mathcal{Z}(m, s, z_{\max})$, if $z + p\mathbb{Z}$ lies in $\mathcal{S}_d(p)$ for all $p \mid b$, check whether $\Delta(d, z)$ is square, and, if so, output the pair (d, z).

Remark 3.6: The following remarks apply to the implementation of *Algorithm 3.5*.

• The algorithm can be easily parallelized by restricting the range of p_1 and, for very small values of p_1, fixing p_1 and restricting the range of p_2.

• The recursive enumeration of $d_0 = p_1^{e_n} \cdots p_n^{e_n}$ ensures that, typically, only the value of $p_n^{e_n}$ changes from one d_0 to the next, allowing the product $\mathcal{C}(p_1^{e_1}) \times \cdots \times \mathcal{C}(p_n^{e_n})$ to be updated incrementally rather than recomputed for each d_0.

• The sets $\mathcal{C}(p^e)$ are precomputed for $p \le \sqrt{d_{\max}}$, as are the sets $\mathcal{A}_d(q)$ for each $d \in \{1, \dots q-1\}$ not divisible by 3, and the sets $\mathcal{S}_d(p)$ for each $p \in A$ and $d \in \{1, \dots, p-1\}$. This allows the sets $\mathcal{Z}(m)$ to be efficiently enumerated using an explicit form of the Chinese remainder theorem that requires very little space. We shall refer to this procedure as CRT enumeration.

• For $p \in A$, the precomputed sets $\mathcal{S}_d(p)$ for $d \in \{1, \dots, p-1\}$ are also stored as bitmaps, as are Cartesian products of pairs of these sets and certain triples; this facilitates testing whether $z + p\mathbb{Z}$ lies in $\mathcal{S}_d(p)$ for $p \mid b$.

Example 3.7: For $k = 33$ and $d = 5$, we have $\mathcal{C}(d) = \{2\}$ and $\operatorname{sgn} z = +1$. For $z_{\max} = 10^{16}$, this leaves 2×10^{15} candidate pairs $(5, z)$ to check. We have $\#\mathcal{A}_d(q) = 14$ with $q = 891$, which reduces this to approximately 3.143×10^{13} candidate pairs. The table below shows the benefit of including additional primes $p \mid a$.

$p \mid a$	$\#\mathcal{S}_d(p)$	$\#\mathcal{Z}(m)$	m	$\#\mathcal{Z}(m, s, z_{\max})$
—	—	14	4,455	3.143×10^{13}
2	1	14	8,910	1.571×10^{13}
7	1	14	62,370	2.245×10^{12}
13	3	42	810,810	5.180×10^{11}
17	9	378	13,783,770	2.742×10^{11}
23	12	4,536	317,026,710	1.431×10^{11}
29	15	68,040	9,193,774,590	7.401×10^{10}
43	19	1,292,760	395,332,307,370	3.270×10^{10}
67	27	34,904,520	26,487,264,593,790	1.318×10^{10}
103	43	1,500,894,360	2,728,188,253,160,370	5.501×10^{9}

The net gain is a factor of more than 363,541 over the naïve approach; we gain a factor of about 63 from cubic reciprocity and local constraints mod q, and a factor of about 5,712 from the $p \mid a$. In general, including auxiliary $p \mid a$ ensures that the number of (d, z) we need to consider for small values of d is a negligible proportion of the total computation.

Remark 3.8: With CRT enumeration, we avoid the need to store the sets $\mathcal{Z}(m)$, analogs of which were explicitly constructed in ref. 4. This greatly reduces the memory required when d is small. In this way, we no longer rely on computations of integral points on the elliptic curve defined by Eq. **1.2** to rule out very small values of d. Nevertheless, we note that one can improve the integral point search carried out in ref. 4, using a trick of Bremner (13) to pass to a 3-isogenous curve. Using this approach, we were able to unconditionally rule out any solutions to Eq. **1.1** with $d \le 100$ for the k listed in Eq. **1.3**, and with $d \le 20,000$ assuming the GRH. It is thus now possible to certify, under GRH, Bremner's heuristic search of the same region in 1995.

4. Heuristics

In this section, we present a heuristic analysis of the distribution of solutions to Eq. **1.1** for a fixed k. We then use this to optimize the choice of the ratio $R := z_{\max}/d_{\max}$.

From Eq. **2.1**, we see that, on $V = \{(x, y, z) \in \mathbb{R}^3 : x^3 + y^3 + z^3 = 0,\ |x| \ge |y| \ge |z|\}$, the proportion of the real density contributed by points satisfying $y/z \in [t_1, t_2]$ is

$$4\sigma_\infty^{-1} \int_{t_1}^{t_2} \frac{dt}{(t^3+1)^{2/3}}. \qquad \textbf{[4.1]}$$

Given a large solution $(x, y, z) \in \mathbb{Z}^3$ to $x^3 + y^3 + z^3 = k$, with $|x| \ge |y| \ge |z|$, the projective point $[x : y : z] \in \mathbb{P}^2(\mathbb{R})$ lies close to the Fermat curve $x^3 + y^3 + z^3 = 0$. We conjecture that, for fixed k, the ratios y/z are distributed as above: The proportion of points (ordered by any height function as in section A) with $y/z \in [t_1, t_2]$ should converge to the quantity in Eq. **4.1**.

Let us assume that this is the case and work out the distribution of $r := -\frac{z}{x+y}$ for $(x, y, z) \in V$. We have

$$-\frac{y}{x} = \frac{2r^3 + 1 - \sqrt{12r^3 - 3}}{2(r^3 - 1)} \quad \text{and} \quad -\frac{z}{x} = \frac{r(\sqrt{12r^3 - 3} - 3)}{2(r^3 - 1)},$$

so that

$$t := \frac{y}{z} = \frac{\sqrt{12r^3 - 3} - 3}{6r} \quad \text{and} \quad (t^3 + 1)^{-2/3} \frac{dt}{dr} = \sqrt{\frac{3}{4r^3 - 1}}.$$

Hence, for any $R \ge \alpha^{-1}$, we have

$$\Pr[r \le R] = 1 - \Pr[r > R] = 1 - 4\sigma_\infty^{-1} \int_R^\infty \sqrt{\frac{3}{4r^3 - 1}}\, dr = 1 - cK(R),$$

where $c = 4\sqrt{3}\sigma_\infty^{-1} = 6\sqrt{3}\frac{\Gamma(2/3)}{\Gamma(1/3)^2} = 1.96084321968938583\ldots$ and

$$K(R) := \int_R^\infty \frac{dr}{\sqrt{4r^3-1}} = R^{-1/2}\sum_{j=0}^{\infty}\frac{\binom{j-\frac{1}{2}}{j}}{1+6j}(4R^3)^{-j}.$$

Thus, the values of $1 - cK(r)$ should be uniformly distributed on $[0, 1]$. To test this hypothesis, we plotted the cumulative distribution of $1 - cK(-z/(x+y))$ over the points of the Huisman dataset with $10^{7.5} < |x| \le 10^{15}$ versus that of a uniform random variable; see Fig. 2.

Example 4.1: For our solution to $x^3 + y^3 + z^3 = 3$, we have

$$r \approx 4.36 \times 10^6 \quad \text{and} \quad cK(r) \approx 9.39 \times 10^{-4},$$

so this solution was an approximately 1-in-1,000 event. This is also reflected by the fact that the solution is highly skewed, with $|x|$ and $|y|$ both much larger than $|z|$.

We use this analysis to optimize the choice of $R = z_{\max}/d_{\max}$ as follows. We assume that a given divisor $d \in \mathbb{Z}_{>0}$ occurs with probability κ_d/d, where κ_d is an arithmetic factor (depending on k) encoding the local solubility, in such a way that

$$\sum_{d\le x}\kappa_d = \rho x + O(x/\log^2 x), \quad \text{for some constant } \rho > 0.$$

By partial summation, it follows that there exists C such that

$$\sum_{d\le x}\frac{\kappa_d}{d} = \rho\log x + C + o(1) \quad \text{and} \quad \sum_{d\le x}\kappa_d f(d) = (\rho + o(1))\int_0^x f(u)\,du \quad \text{as } x\to\infty,$$

for any monotonically decreasing function f satisfying $f(u) \asymp u^{-s}$ for some $s \in (0, 1)$. In turn, we expect to find z in a fixed arithmetic progression modulo $d \le d_{\max}$ with probability

$$\Pr[|z| \le z_{\max} \mid d \text{ fixed}] = \Pr[r \le z_{\max}/d] = 1 - cK(z_{\max}/d).$$

Hence, the number of solutions that we expect to find is given by

$$\begin{aligned}\sum_{d\le d_{\max}}\frac{\kappa_d}{d}\left(1 - cK\left(\frac{z_{\max}}{d}\right)\right) &= \rho\log d_{\max} + C + o(1) - (\rho + o(1))c\int_0^{d_{\max}} K\left(\frac{z_{\max}}{u}\right)\frac{du}{u}\\ &= \rho\log d_{\max} + C - \rho c\int_{z_{\max}/d_{\max}}^{\infty} K(r)\frac{dr}{r} + o(1).\end{aligned}$$

Taking $d_{\max} = \alpha z_{\max}$ recovers Heath-Brown's conjecture, provided that $\rho = \rho_{\text{sol}}$.

Next, suppose that the total running time is $T(d_{\max}, z_{\max})$, and let T_d and T_z denote its partial derivatives. Let $d_{\max}$ be defined implicitly in terms of $R = z_{\max}/d_{\max}$ so that $(d_{\max}, z_{\max})$ remains on a level set for T, meaning that

$$T(d_{\max}, Rd_{\max}) = \text{constant}.$$

Differentiating with respect to R, we have

$$T_d(d_{\max}, Rd_{\max})\frac{\partial d_{\max}}{\partial R} + T_z(d_{\max}, Rd_{\max})\left(d_{\max} + R\frac{\partial d_{\max}}{\partial R}\right) = 0.$$

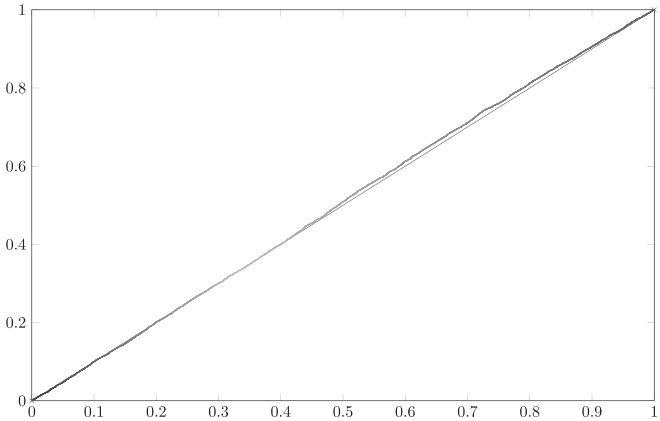

Fig. 2. Cumulative distribution of $1 - cK(-z/(x+y))$ over solutions (x, y, z) in the Huisman dataset with $\max\{|x|, |y|, |z|\} \in [10^{7.5}, 10^{15}]$, versus a uniform random variable.

We seek to maximize the expected solution count, which, to leading order, is

$$\rho \log d_{\max} + C - \rho c \int_R^\infty K(r) \frac{dr}{r}.$$

Differentiating with respect to R, this gives

$$\frac{\rho}{d_{\max}} \frac{\partial d_{\max}}{\partial R} + \frac{\rho c K(R)}{R} = 0,$$

so that $\frac{\partial d_{\max}}{\partial R} = -c d_{\max} K(R)/R$. Substituting this into the above, we obtain

$$\frac{T_d(d_{\max}, R d_{\max})}{T_z(d_{\max}, R d_{\max})} = R\left(\frac{1}{cK(R)} - 1\right) \approx c^{-1} R^{3/2}\left(1 - \frac{1}{56R^3}\right) - R =: C_R.$$

In Table 3, we show computed T_d/T_z ratios for $k = 3$ and various values of R and $d_{\max}$. For a given $d_{\max}$, we wish to choose R so that $T_d/T_z \approx C_R$. It is difficult to measure T_d/T_z precisely; it is the ratio of two small numbers, and this ratio is easily influenced by small differences in timings when running computations on different hardware. To compute the values below, we used a single hardware platform and took medians of five runs to compute each row.

From Table 3, we can see that, for $k = 3$ and $d_{\max} \geq 2^{35}$, the optimal choice of R is greater than 32, and, for $d_{\max} \geq 2^{50}$, it is greater than 64. For other values of k, the pattern is similar, although the T_d/T_z vary slightly; this is to be expected, given the varying benefit of cubic reciprocity constraints.

5. Computational Results

A. Implementation. We implemented the algorithm described in section A using the gcc C compiler (14) and the primesieve library for fast prime enumeration (15). We parallelized by partitioning the set of primes $p \leq d_{\max}$ into subintervals $[p_{\min}, p_{\max}]$ of suitable size, with the work distributed across jobs that checked all of the (d, z) candidates with the largest prime factor $p_1 \mid d$ lying in the assigned interval. Each job was run on a separate machine, with local parallelism achieved by distributing the p_1 across available cores (and, for small values of p_1, also distributing the p_2), as noted in *Remark 3.6*. When choosing the number of jobs and the sizes of the intervals $[p_{\min}, p_{\max}]$, we use the ρ_{ap} density estimates derived in section A, as noted in *Remark 2.2*.

We used a standard Tonelli–Shanks approach to computing cube roots modulo primes; this involves computing a discrete logarithm in the 3-Sylow subgroup of $(\mathbb{Z}/p\mathbb{Z})^\times$, using $O(1)$ group operations on average, and $O(1)$ exponentiations. Hensel lifting was used to compute cube roots modulo prime powers; these were precomputed and cached for all prime powers up to $\min\{p_{\max}, \sqrt{d_{\max}}\}$. For the values of $d_{\max}$ that we used, this precomputation typically takes just a few seconds, and the cache size is well under 1 gigabyte. We use Montgomery representation (16) for performing arithmetic in $(\mathbb{Z}/p^r\mathbb{Z})^\times$, but switch to standard integer representation and use Barrett reduction (17) during CRT enumeration of cube roots of k modulo d, and when sieving arithmetic progressions via auxiliary primes.

For the k of interest, the sets $\mathcal{A}_d(q)$ giving constraints modulo the integer q defined in *Lemma 3.3* for admissible pairs (d, z) were precomputed and cached; again, this takes only a few seconds for the largest values of k. In order to avoid using arithmetic progressions of modulus larger than $z_{\max}$, we project these constraints to residue classes modulo a suitably chosen divisor of q when $qd > z_{\max}$.

B. Computations. In September 2019, we ran computations for the 11 unresolved $k \leq 1{,}000$ listed in Eq. **1.3** on Charity Engine's crowd-sourced compute grid consisting of approximately 500,000 personal computers. For this initial search, we used $z_{\max} = 10^{17}$ and $d_{\max} = \alpha z_{\max}$ to search for all solutions to Eq. **1.1** with $\min\{|x|, |y|, |z|\} \leq 10^{17}$. This search yielded the solutions for $k = 42$, $k = 165$, and $k = 906$ listed in the Introduction. We then ran a search for $k = 3$ using $z_{\max} = 10^{18}$ and $d_{\max} = \alpha z_{\max}/9$ and found the solution for $k = 3$ listed in the Introduction. These computations involved a total of several hundred core-years but were completed in just a few weeks (it is difficult to give more precise estimates of the computational costs, due to variations in processor speeds and resource

Table 3. T_d/T_z versus C_R for various values of $d_{\max}$ and $R = z_{\max}/d_{\max}$ for $k = 3$

R	$d_{\max}$	$z_{\max}$	T_d	T_z	T_d/T_z	C_R
32	2^{35}	2^{40}	2.804×10^{-08}	2.359×10^{-10}	118.9	60.3
32	2^{40}	2^{45}	2.738×10^{-08}	2.247×10^{-10}	121.8	60.3
32	2^{45}	2^{50}	2.922×10^{-08}	2.175×10^{-10}	134.4	60.3
32	2^{50}	2^{55}	3.113×10^{-08}	2.150×10^{-10}	144.8	60.3
32	2^{55}	2^{60}	3.678×10^{-08}	2.100×10^{-10}	175.2	60.3
64	2^{35}	2^{41}	3.140×10^{-08}	1.813×10^{-10}	173.2	197.1
64	2^{40}	2^{46}	2.771×10^{-08}	1.730×10^{-10}	160.2	197.1
64	2^{45}	2^{51}	3.112×10^{-08}	1.613×10^{-10}	192.9	197.1
64	2^{50}	2^{56}	3.187×10^{-08}	1.506×10^{-10}	211.6	197.1
64	2^{55}	2^{61}	3.862×10^{-08}	1.612×10^{-10}	239.6	197.1
128	2^{35}	2^{42}	3.749×10^{-08}	1.238×10^{-10}	302.8	618.5
128	2^{40}	2^{47}	3.407×10^{-08}	1.216×10^{-10}	280.2	618.5
128	2^{45}	2^{52}	3.826×10^{-08}	1.530×10^{-10}	250.1	618.5
128	2^{50}	2^{57}	3.768×10^{-08}	1.185×10^{-10}	318.0	618.5
128	2^{55}	2^{62}	4.096×10^{-08}	1.091×10^{-10}	375.4	618.5

availability in a crowd-sourced computation). Subsequently, over the course of 2020, Charity Engine conducted a search at lower priority for the remaining eight candidate values of k, with $z_{\max} = 10^{19}$ and $d_{\max} = z_{\max}/54$; this yielded the solution for $k = 579$ in January 2021.

Remark 5.1: While, in principle, these searches rule out the existence of any solutions that were not found, we are reluctant to make any unconditional claims. Despite putting in place measures to detect failures, including counting the primes that were enumerated (these counts can be efficiently verified after the fact), there is always the possibility of undetected hardware or software errors, especially on a large network of personal computers that typically do not have error correcting memory.

In order to verify the minimality of the solution we found for $k = 3$, we ran a separate verification with $z_{\max}$ equal to 472,715,493,453,327,032, the absolute value of the z in our solution, and $d_{\max} = \alpha z_{\max}$. This search was run on Google's Compute Engine (19) and found no solutions other than those already known. These computations were run on 8-core (16 virtual CPU) instances equipped with Intel Xeon processors in the Sandybridge, Haswell, and Broadwell families running at 2.0 GHz or 2.2 GHz. Using 155,579 nodes, the computation took less than 4 h and used approximately 120 core-years. We detected errors in 5 of the 155,579 runs, which were corrected upon rerunning the computations. Barring the existence of any undetected errors, these computations rule out any smaller solutions for $k = 3$ other than those we now know.

To assess the benefit of the theoretical and algorithmic improvements introduced here, we searched for solutions to $k = 33$ using $R = 64$, which is close to the optimal choice for $d_{\max}$ in the range $[2^{40}, 2^{50}]$. The general search strategy we envision is to start with a value of $d_{\max}$ for which all solutions with $|z| \leq R d_{\max}$ are known, where R is chosen optimally for $d_{\max}$. One would then successively double $d_{\max}$, adjusting R as necessary, and run a search using $z_{\max} = R d_{\max}$. If one takes care to avoid checking the same admissible (d, z) twice, the total time is approximately equal to a single complete search using the final values of $d_{\max}$ and R (one expects R to be increasing). The first $d_{\max} = 2^n$ sufficient to find a solution for $k = 33$ with this strategy is $d_{\max} = 2^{47}$, for which we choose $R = 64$, yielding $z_{\max} = 2^{53}$. Using 2.8-GHz Intel processors in the Skylake family, this search finds the known solution for $k = 33$ in 107 core-days. The search in ref. 4 using $z_{\max} = 10^{16}$ and $d_{\max} = \alpha z_{\max}$ took 3,145 core-days running mostly on 2.6-GHz Intel processors in the Sandybridge family. After adjusting for the difference in processor speeds and $z_{\max}$ values, our approach finds the solution for $k = 33$ approximately 25 times faster.

In the future, we hope to use this strategy to search for solutions for the seven $k \leq 1{,}000$ that remain unresolved:

$$114,\ 390,\ 627,\ 633,\ 732,\ 921,\ 975. \qquad \textbf{[5.1]}$$

Data Availability. Code has been deposited in GitHub at https://github.com/AndrewVSutherland/SumsOfThreeCubes (19).

ACKNOWLEDGMENTS. We are extremely grateful to Charity Engine for providing the computational resources used for this project, and, in particular, to Mark McAndrew, Matthew Blumberg, and Rytis Slatkevičius, who were responsible for running these computations on the Charity Engine compute grid. We thank Roger Heath-Brown for several stimulating discussions; in particular, his conversation with A.R.B. in the Nettle and Rye on February 27, 2019 informed the analysis presented in section 4. A.V.S. also acknowledges the support of the Simons Foundation (Award 550033).

1. D. R. Heath-Brown, The density of zeros of forms for which weak approximation fails. *Math. Comp.* **59**, 613–623 (1992).
2. L. J. Mordell, On the integer solutions of the equation $x^2 + y^2 + z^2 + 2xyz = n$. *J. London Math. Soc.* **28**, 500–510 (1953).
3. J. C. P. Miller, M. F. C. Woollett, Solutions of the Diophantine equation $x^3 + y^3 + z^3 = k$. *J. London Math. Soc.* **30**, 101–110 (1955).
4. A. R. Booker, Cracking the problem with 33. *Res. Number Theory* **5**, 26 (2019).
5. D. R. Heath-Brown, W. M. Lioen, H. J. J. te Riele, On solving the Diophantine equation $x^3 + y^3 + z^3 = k$ on a vector computer. *Math. Comp.* **61**, 235–244 (1993).
6. S. G. Huisman, Newer sums of three cubes. arXiv [Preprint] (2016). https://arxiv.org/abs/1604.07746 (Accessed 2 July 2020).
7. R. Harron, The shapes of pure cubic fields. *Proc. Amer. Math. Soc.* **145**, 509–524 (2017).
8. The PARI Group, PARI/GP, 2.11.21 edition. https://pari.math.u-bordeaux.fr/. Accessed 2 July 2020.
9. R. Kato, A remark on the Wiener-Ikehara Tauberian theorem. *Comment. Math. Univ. St. Pauli* **64**, 47–58 (2015).
10. J. W. S. Cassels, A note on the Diophantine equation $x^3 + y^3 + z^3 = 3$. *Math. Comp.* **44**, 265–266 (1985).
11. J. L. Colliot-Thélène, O. Wittenberg, Groupe de Brauer et points entiers de deux familles de surfaces cubiques affines. *Amer. J. Math.* **134**, 1303–1327 (2012).
12. K. Ireland, M. Rosen, *A Classical Introduction to Modern Number Theory* (Graduate Texts in Mathematics, Springer-Verlag, New York, NY, ed. 2, 1990), vol. 84.
13. A. Bremner, "On sums of three cubes in Number theory (Halifax, NS, 1994)" in *CMS Conference Proceedings*, K. Dilcher, Ed. (American Mathematical Society, Providence, RI, 1995), vol. 15, pp. 87–91.
14. The GCC Team, GCC: The GNU compiler collection, 7.4.0 edition. https://gcc.gnu.org/. Accessed 2 July 2020.
15. K. Walisch, Data from "primesieve." GitHub. https://github.com/kimwalisch/primesieve. Accessed 2 July 2020.
16. P. L. Montgomery, Modular multiplication without trial division. *Math. Comp.* **44**, 519–521 (1985).
17. P. Barrett, "Implementing the Rivest Shamir and Adleman public key encryption algorithm on a standard digital signal processor" in *Advances in Cryptology—CRYPTO '86 (Santa Barbara, Calif., 1986)*, A. M. Odlyzko, Ed. (Lecture Notes in Computer Science, Springer, Berlin, Germany, 1987) vol. 263, pp. 311–323.
18. Google LLC, Google compute engine documentation. https://cloud.google.com/compute/docs/. Accessed 2 July 2020.
19. A. R. Booker, A. V. Sutherland, Sums of three cubes. GitHub. https://github.com/AndrewVSutherland/SumsOfThreeCubes. Deposited 15 February 2021.